

- минимизировать влияние субъективных факторов (участие человека) на подготовку и передачу данных;
- обеспечивать репрезентативность, дублировать поступающую информацию из разных источников;
- по возможности увеличить частоту обновления (актуализации) данных.

Литература:

1. Указ Президента Российской Федерации от 7 мая 2018 года № 204 «в редакции 21.07.2020 г. № 474 «О национальных целях развития Российской Федерации на период до 2024 года». – URL: <http://publication.pravo.gov.ru/Document/View/0001201805070038> (дата обращения 11.10.2021).

2. Указ Президента Российской Федерации от 21.07.2020 № 474 «О национальных целях развития Российской Федерации на период до 2030 года». – URL: <http://publication.pravo.gov.ru/Document/View/0001202007210012?index=3&rangeSize=1> (дата обращения 11.10.2021).

3. Бизнес ПРОСТ, Что такое достоверность информации? Описание и определение понятия. – URL: <https://biznesprost.ru/dostovernost-informacii.html> (дата обращения 30.09.2021).

4. *Арбатская Т.Г.* К вопросу о сущности категории «достоверность» // Международный бухгалтерский учет. – 2015. – № 8. – С. 17-32.

5. *Козлов А.Д., Нога Н.Л.* Риски информационной безопасности корпоративных информационных систем при использовании облачных технологий // Управление риском. – 2019. – № 3. – С. 31-46.

Сомов С.К.

Проблема оптимизации схемы восстановления разрушенного оперативного резерва данных в распределенных системах

Аннотация: В работе рассмотрена проблема восстановления разрушенных оперативных данных, используемых в распределенной системе, с помощью специального восстановительного резерва. Предложены различные схемы организации процесса восстановления

разрушенных данных. Предложена формальная модель восстановления данных, на основе которой поставлена задача поиска схемы восстановления разрушенных данных, оптимальной по различным критериям оптимальности решения задачи.

Ключевые слова: распределенные системы, оперативное резервирование, восстановительное резервирование, схема восстановления разрушенного резерва

Для обеспечения высокой степени сохранности информации, используемой в распределенных информационных системах, широко используется метод информационной избыточности. А именно – создание и размещение в узлах компьютерной сети, на основе которой работает РСОД, оперативного резерва, состоящего из некоторого количества копий и/или предысторий массивов данных [1]. Такой подход сильно снижает величину вероятности потери или искажения данных, но не исключает полностью возможность искажения или потери данных, используемых в распределенной системе и распределенных в нескольких узлах системы. Данная проблема может быть эффективно устранена путем использования двух методов восстановительного резервирования [2]:

1) Первый метод заключается в том, что для восстановления разрушенных в узле системы оперативных данных используется восстановительный резерв (ВР), в качестве которого используется неразрушенный оперативный резерв, расположенный в узле системы, ближайшем к узлу с разрушенными данными.

2) Согласно второму методу для восстановления разрушенных в некоем узле системы данных используется специальный резерв из копий и/или предысторий массивов данных – архив магнитных носителей (АМН). Данный архив используется только для восстановления разрушенных данных. Данный метод заключается в том, что в узле с АМН создаются копии массивов данных, которые затем пересылаются в узел с разрушенными данными.

Для восстановления разрушенных данных используется одна из двух стратегий восстановления: В-1 или В-2 [2]. При использовании стратегии В-1 в узле с ВР последовательно, одна за другой, на основе данных ВР создается необходимое количество копий

массива данных ОР, который необходимо восстановить. Восстановленные копии затем пересылаются в узел с разрушенным ОР. Восстановление разрушенного ОР согласно стратегии В-2 происходит аналогично стратегии В-1 за одним исключением: при создании очередной копии массива данных используется не только ВР, но и все восстановленные на данный момент массивы данных.

С учетом того, что РСОД состоит из множества узлов, соединенных между собой каналами связи, для восстановления разрушенных данных в некотором узле системы возможно использование различных схем восстановления.

Предположим, что в некотором i -м узле системы использовался оперативный резерв из m копий массива данных, которые в силу некоторых причин были разрушены. Предположим, что в узле j , ближайшем к узлу i с разрушенным резервом, размещен восстановительный резерв, содержащий копии разрушенных в узле i массивов данных. Тогда восстановительный резерв в узле j можно использовать следующим образом. Восстановить с помощью ВР j -го узла y ($1 \leq y \leq m$) копий разрушенного массива. Затем по каналам связи переслать в i -й узел y восстановленных копий. Остальные $(m - y)$ копий разрушенного массива восстанавливаются простым копированием в самом узле j . В итоге возникает задача определения такого количества y восстанавливаемых в узле с ВР копий массива, чтобы достигалось оптимальное значение критерия оптимальности задачи.

На примере стратегии В-1 сформулируем задачу поиска оптимальной схемы восстановления разрушенных данных. В качестве критериев оптимизации можно использовать следующие три варианта:

- минимум средних затрат $Z(y)$ системы на восстановление разрушенного ОР;
- минимум среднего времени $E(y)$ затраченного на восстановление разрушенного ОР;
- максимум вероятности $P(y)$ восстановления разрушенного оперативного резерва.

Предположим, что в узле с ВР копии массивов данных восстанавливаются последовательно и пересылаются в узел с разрушенным ОР также последовательно. В этом случае на восстановление разрушенного резерва будет затрачено время,

величина которого будет состоять из трех компонент: время, затраченное в узле с ВР на получение y копий массива данных, затраты времени на передачу полученных копий в узел с разрушенным ОР по каналам связи и время, затраченное в узле с ОР на создание в нем остальных $(m-y)$ копий массива.

Предположим, что на восстановление одной копии массива в узле с ВР затрачивается T единиц времени. Обозначим через t среднее время передачи по каналам связи одной копии массива данных. Тогда для оценки среднего времени $E(y)$ восстановления y копий массива данных требуется рассмотреть следующие варианты соотношений времени t и времени T , представленные на рисунке 1.

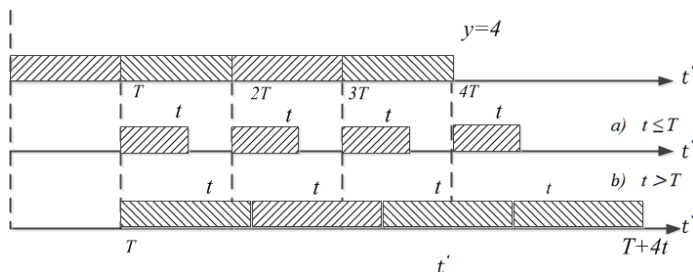


Рисунок 1 – Соотношения между значениями времен T и t

а) $t \leq T$ – ситуация при которой среднее время t передачи по каналам связи между двумя узлами одной копии массива данных не превышает среднего времени создания копии такого же массива данных в узле с ВР.

б) $t > T$ – в этой ситуации среднее время t передачи по каналам связи одной копии массива больше, чем время ее создания в узле с восстановительным резервом.

С учетом представленных выше вариантов среднее время $E(y)$ восстановления оперативного резерва, разрушенного в узле системы, будет равно:

$$E(y) = \begin{cases} yT + t + (m - y)\tau & \text{при } t \leq T \\ yt + T + (m - y)\tau & \text{при } t > T \end{cases} \quad (1)$$

В формуле (1) τ это величина среднего времени создания в узле с ОР одной копии массива данных.

Разрушенный в узле системы оперативный резерв будет успешно восстановлен с вероятностью:

$$P(y) = \beta^m (\rho r \beta^{-1}) \quad (2)$$

В формуле (2):

– ρ – это вероятность, с которой в узле с ВР создается копия массива данных, необходимая для восстановления разрушенного ОР;

– r – это вероятность успешной передачи восстановленной в узле с ВР копии массива данных в узел, в котором необходимо восстановить разрушенный ОР;

– β – это вероятность, с которой в узле с ОР создается копия массива данных, необходимая для полного восстановления разрушенного ОР.

На восстановление разрушенного ОР в узле системы в среднем тратятся ресурсы $Z(y)$, которые равны:

$$Z(y) = y(D + d^* - H) + H * m \quad (3)$$

Здесь: H – затраты ресурсов системы на создание одной копии массива в узле с ОР, d^* – стоимость использования каналов связи для передачи между двумя узлами одной копии восстанавливаемого массива, D – затраты ресурсов системы на восстановления одной копии массива в узле с ВР.

Предположим, что выполняется соотношение: $M \leq y \leq N$ ($1 \leq M, N \leq m$) и выполним анализ значений для $E(y)$. В соответствии с формулой (1) получаем, что $E(y) < 0$ при $t \leq T, T < \tau$ и $t > T, t < \tau$. Из этого утверждения следует, что:

$$\begin{aligned} \text{при } t \leq T \quad \min E(y) &= \begin{cases} m\tau + t + N(T - \tau) & \text{при } T < \tau \\ m\tau + t + M(T - \tau) & \text{при } T \gg \tau \end{cases} \\ \text{при } t > T \quad \min E(y) &= \begin{cases} m\tau + T + N(t - \tau) & \text{при } t < \tau \\ m\tau + T + M(t - \tau) & \text{при } t \gg \tau \end{cases} \end{aligned} \quad (4)$$

Для $P(y)$ и $Z(y)$ получим следующие выражения:

$$\begin{aligned} \max P(y) &= \begin{cases} P(N) & \text{при } \rho r > \beta \\ P(M) & \text{при } \rho r \leq \beta \end{cases} \\ \max Z(y) &= \begin{cases} Z(N) & \text{при } (D + d^*) < H \\ Z(M) & \text{при } (D + d^*) \geq H \end{cases} \end{aligned} \quad (5)$$

С учетом полученных результатов задача определения оптимальной схемы восстановления разрушенного ОР с

использованием в качестве критерия минимума затрат системы будет иметь следующую формулировку:

$$Z(y) \rightarrow \min \quad (6)$$

при следующих ограничениях: $P(y) \geq \bar{P}$; $E(y) \leq \bar{E}$; $y \in \{1, 2, \dots, m\}$.

Здесь \bar{P} и \bar{E} это ограничения на минимальную вероятность успешного восстановления и на максимальные затраты системы на восстановление разрушенного ОР, соответственно.

Аналогичным путем формулируются задачи поиска наилучшей схемы восстановления оперативного резерва, разрушенного в некотором узле системы, с помощью восстановительного резерва с использованием в качестве критерия оптимальности задачи минимума среднего времени восстановления разрушенного ОР и максимума вероятности восстановления разрушенного ОР.

Заключение

В работе предложена формальная модель возможных схем восстановления разрушенного оперативного резерва в распределенной системе. На ее основе предложены формулировки задач оптимизации схемы восстановления разрушенных данных с использованием нескольких критериев оптимальности решения данных задач.

Работа выполнена в рамках темы: «Фундаментальные исследования по направлению «Модели, методы анализа и синтеза структуры и сценариев развития социально-экономических и технических систем управления, повышения их управляемости и безопасности функционирования в условиях неопределенности, структурных возмущений и чрезвычайных ситуаций» № 0052-2019-0011

Литература:

1. *Сомов С.К.* Сохранность информации в распределенных системах обработки данных. – М.: ИПУ РАН, 2019. – 254 с.
2. *Микрин Е.А., Сомов С.К.* Анализ эффективности стратегий восстановления информации в распределенных системах обработки данных // Информационные технологии и вычислительные системы. – 2016. – №3. – С. 5-19.